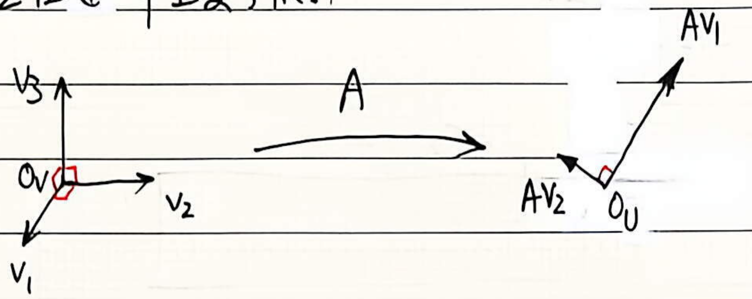
5. SVD (singular value decomposition)

(1) Motivation:

① let $A: \mathbb{C}^n \to \mathbb{C}^m$，我们希望在 $\mathbb{C}^n$ 中找到一组 orthonormal basis，其通过 A 作用后，仍然在 $\mathbb{C}^m$ 中正交；i.e.



or written:

$$v_i \perp v_j \implies Av_i \perp Av_j \quad \forall i, j.$$

NOTE. i) A 不一定 injective，因此可能存在映射到 O_U 的可能：

maybe $\exists v_i \in \mathbb{C}^n$, $Av_i = O_U$.

ii) 并不要求 $\{Av_i\}$ 归一化（只需正交），当然我们可以通过归一化来实现，即：

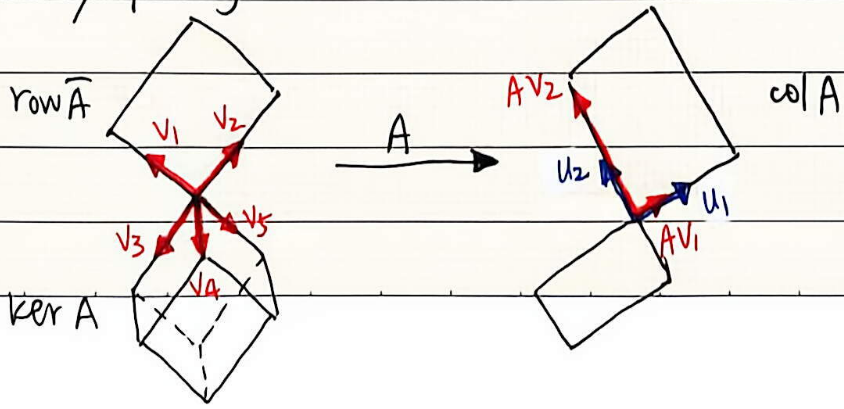
$\exists$ orthonormal basis $\{u_i\}$ that $\text{span}\{u_i\} = \text{col}\,A$, s.t.

↳ 不是 $\mathbb{C}^m$!

$$\text{in row}\,\bar{A} - \left\{ \begin{array}{l} Av_1 = \sigma_1 u_1 \\ Av_2 = \sigma_2 u_2 \\ \quad \vdots \\ \text{"rank"} \quad Av_r = \sigma_r u_r \end{array} \right\} - \text{in col}\,A$$

$$\text{in ker}\,A - \left\{ \begin{array}{l} Av_{r+1} = O_U \\ \quad \vdots \\ Av_n = O_U \end{array} \right\} - \text{at } O_U$$

or visually speaking:



iii) 注意对比这个过程与 eigen decomposition.

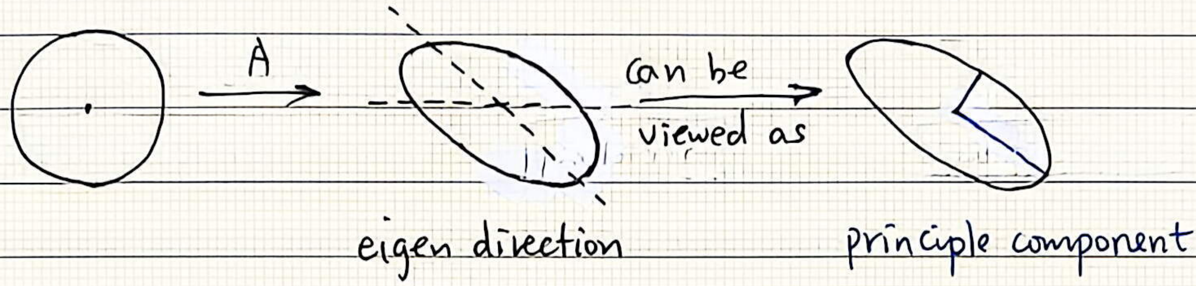
SVD 可以是任意两个向量空间之间线性变换的分解.(维度可以不同!)

② 对于 $\forall A: \mathbb{C}^n \to \mathbb{C}^m$, 一定能找到 $\{v_i\}$ 满足上述条件吗?

一定!

Lemma*. Circles must be turned into ellipses under linear transformation.

E.g.

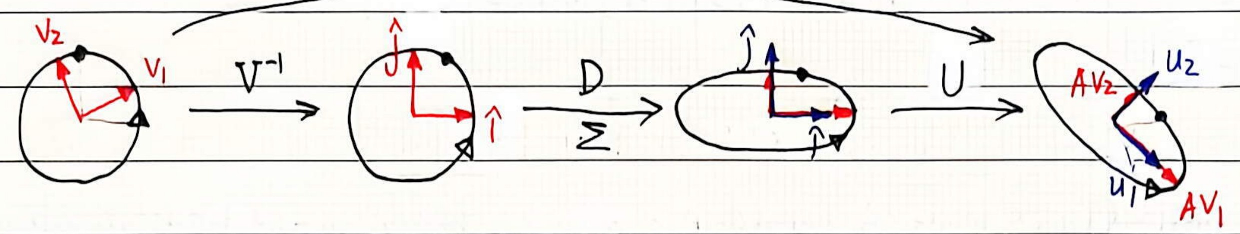


只要你承认所有的圆在线性变换下都变成椭圆,则这个线性变换一定可以分解为"旋转+拉伸+旋转"的复合!

(2) Understanding SVD.

首先考虑2维同构 $A\in \mathrm{Aut}_{v.s.}(\mathbb{R}^2)$，以 marked unit circle（"△"，"•"）在变换 A 下的像即可写出 A 的 SVD 分解：

$$A=\begin{pmatrix}\cdot&\cdot\\ \cdot&\cdot\end{pmatrix}$$



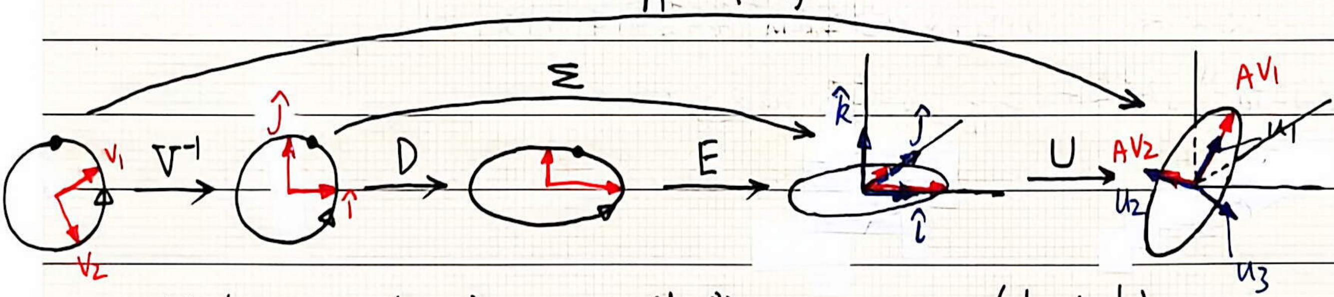
$$V=\begin{pmatrix}|&|\\ v_1&v_2\\ |&|\end{pmatrix}\qquad D=\Sigma=\begin{pmatrix}\sigma_1&\\&\sigma_2\end{pmatrix}\qquad U=\begin{pmatrix}|&|\\ u_1&u_2\\ |&|\end{pmatrix}$$

and $V^{-1}=V^T$ and $U^{-1}=U^T$

$\boxed{\text{change of basis}}$ $\boxed{\text{stretching}}$ $\boxed{\text{linear transformation}}$

再考虑 $A\in \mathrm{Hom}_{v.s}(\mathbb{R}^2,\mathbb{R}^3)$：

$$A=\begin{pmatrix}\cdot&\cdot\\ \cdot&\cdot\\ \cdot&\cdot\end{pmatrix}$$



$$V=\begin{pmatrix}|&|\\ v_1&v_2\\ |&|\end{pmatrix}\quad D=\begin{pmatrix}\sigma_1&\\&\sigma_2\end{pmatrix}\quad E=\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}\quad U=\begin{pmatrix}|&|&|\\ u_1&u_2&u_3\\ |&|&|\end{pmatrix}$$

$\boxed{\text{Change of basis}}$ $\boxed{\text{stretching}}$ $\boxed{\text{Dimension eraser}}$/Adder $\boxed{\text{linear transformation}}$

我们有：

$$A=UEDV^{-1}=U\Sigma V^T=\begin{pmatrix}|&|&|\\ u_1&u_2&u_3\\ |&|&|\end{pmatrix}\begin{pmatrix}\sigma_1&\\&\sigma_2\\ & \end{pmatrix}\begin{pmatrix}|&|\\ v_1&v_2\\ |&|\end{pmatrix}$$

3×3 $\quad 3\times2$ $\quad 2\times2$

NOTE. i) SVD也可以写成 $A=U\Sigma V^{-1}$

之所以写成转置形式是为了强调 $V\in U(n)$

因此一般sandwich term $A=BCD^H$ 都默认 $D\in U(n)$, C为Σ, $B\in U(n)$, 即:

A可以被两个unitary matrix分解!

ii) Stretching的倍数 σ_1, σ_2 称为A的singular value.

一般从左上到右下按照从大到小的顺序排列. (与PCA和低秩分解有关)

e.g.

$$\Sigma_1=\begin{pmatrix} 9.2 & 0 & 0 \\ 0 & 1.3 & 0 \\ 0 & 0 & 0.6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \Sigma_2=\begin{pmatrix} 10.2 & 0 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 & 0 \\ 0 & 0 & 1.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

iii) By convention, $A\in\mathbb{C}^{m\times n}$ 的singular value为正数!

(A的singular value一定为实数 in the first place).

iv) SVD不唯一.

(3) Computing V, Σ and U.

我们将P18. ii)中的式子用矩阵表达 (E.g. $A: \mathbb{C}^{5\times 3}$)

$$\begin{pmatrix} \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}\begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} | & | & | & | & | \\ u_1 & u_2 & u_3 & u_4 & u_5 \\ | & | & | & | & | \end{pmatrix}\begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \\ & & \\ & & \end{pmatrix}$$

or:

$$AV = U\Sigma$$

$$\Rightarrow \quad A = U\Sigma V^H$$

The clever way is to consider AA^H and A^HA:

$$AA^H = (U\Sigma V^H)(U\Sigma V^H)^H = U\Sigma \underbrace{\boxed{V^HV}}_{=I} \Sigma^H U^H = U\Sigma\Sigma^H U^H$$

$$A^HA = (U\Sigma V^H)^H (U\Sigma V^H) = V\Sigma^H \underbrace{\boxed{U^HU}}_{=I} \Sigma V^H = V\Sigma^H\Sigma V^H$$

由于 AA^H 和 A^HA 都为 Hermitian，我们将这个问题化归到了一个 eigen decomp. 的问题，即：

☆ $\begin{cases} A^HA \text{ 的归一化特征向量构成了 } V \\ AA^H \text{ 的归一化特征向量构成了 } U \end{cases}$

奇异值怎么求呢？

从 AA^H 或 A^HA 的特征值中都可以求得 A 的奇异值！

$$\Sigma\Sigma^H = \begin{pmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \sigma_3^2 & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix} \qquad \Sigma^H\Sigma = \begin{pmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2 \end{pmatrix}$$

可见：

$\boxed{A^H A \text{ 或 } AA^H \text{ 的(前) } n \text{ 个特征值的 square root 即为 } A \text{ 的 singular value.}}$